```
                 9        18        27        36        45        54
  5' GCC CTA CCG TGT GCG CAG AAA GAG GAG GCG CTT GCC TTC AGC TTG TGG GAA ATC
                                                                              I

                63        72        81        90        99       108
     CCG AAG ATG GCC AAA GAC AAC TCA ACT GTT CGT TGC TTC CAG GGC CTG CTG ATT
            M   A   K   D   N   S   T   V   R   C   F   Q   G   L   L   I

               117       126       135       144       153       162
     TTT GGA AAT GTG ATT ATT GGT TGT TGC GGC ATT GCC CTG ACT GCG GAG TGC ATC
      F   G   N   V   I   I   G   C   C   G   I   A   L   T   A   E   C   I

               171       180       189       198       207       216
     TTC TTT GTA TCT GAC CAA CAC AGC CTC TAC CCA CTT GAA CTG CTG ACC GAC AAC
      F   F   V   S   D   Q   H   S   L   Y   P   L   E   L   L   T   D   N

               225       234       243       252       261       270
     GAT GAC ATC TAT GGG GCT GCC TGG ATC GGC ATA TTT GTG GGC ATC TGC CTC TTC
      D   D   I   Y   G   A   A   W   I   G   I   F   V   G   I   C   L   F

               279       288       297       306       315       324
     TGC CTG TCT GTT CTA GGC ATT GTA GTG ATC ATG AAG TCC AGC AGG AAA ATT CTT
      C   L   S   V   L   G   I   V   V   I   M   K   S   S   R   K   I   L

               333       342       351       360       369       378
     CTG GCG TAT TTC ATT CTG ATG TTT ATA GTA TAT GCC TTT GAA GTG GCA TCT TGT
      L   A   Y   F   I   L   M   F   I   V   Y   A   F   E   V   A   S   C
```

FIGURE 1A

```
      387         396         405         414         423         432
ATC ACA GCA ACA CAA CGA GAC TTT TTC ACA CCC AAC CTC TTC CTG AAG CAG
 I   T   A   T   Q   R   D   F   F   T   P   N   L   F   L   K   Q

      441         450         459         468         477         486
ATG CTA GAG AGG TAC CAA AAC AGC CCT CCA AAT GAT GAC CAG TGG AAA
 M   L   E   R   Y   Q   N   S   P   P   N   D   D   Q   W   K

      495         504         513         522         531         540
AAC AAT GGA GTC ACC AAA ACC TGG GAC AGG CTC ATG CTC CAG GAC AAT TGC TGT
 N   N   G   V   T   K   T   W   D   R   L   M   L   Q   D   N   C   C

      549         558         567         576         585         594
GGC GTA AAT GGT CCA TCA GAC TGG CAA AAA TAC TCT GCC TTC CGG ACT GAG
 G   V   N   G   P   S   D   W   Q   K   Y   S   A   F   R   T   E

      603         612         621         630         639         648
AAT GAT GCT GAC TAT CCC TGG CCT CGT CAA TGC TGT GTT ATG AAC AAT CTT
 N   D   A   D   Y   P   W   P   R   Q   C   C   V   M   N   N   L

      657         666         675         684         693         702
AAA GAA CCT CTC AAC CTG GAG GCT TGT AAA CTA GGC GTG CCT GGT TTT TAT CAC
 K   E   P   L   N   L   E   A   C   K   L   G   V   P   G   F   Y   H

      711         720         729         738         747         756
AAT CAG GGC TGC TAT GAA CTG ATC TCT GGT CCA ATG AAC CGA CAC GCC TGG GGG
 N   Q   G   C   Y   E   L   I   S   G   P   M   N   R   H   A   W   G
```

FIGURE 1B

```
        765       774       783       792       801       810
        GTT GCC TGG TTT GGA TTT GCC ATT CTC TGC TGG ACT TTT TGG GTT CTC CTG GGT
         V   A   W   F   G   F   A   I   L   C   W   T   F   W   V   L   L   G

        819       828       837       846       855       864
        ACC ATG TTC TAC TGG AGC AGA ATT GAA TAT TAA AGT GTT GCC ACC ATA
         T   M   F   Y   W   S   R   I   E   Y   *

        873       882       891       900       909       918
        CCT CCT CCG AGT GAC TCT GGA TTT GGT GCT GGA ACC AGC TCT CTC CTA ATA

        927       936       945       954       963       972
        TTC CAC GTT TGT GCC CCA CAC TAA CGT GTG TGT CTT ACA TTG CCA AGT CAG ATG

        981       990       999       1008      1017      1026
        GTA CGG ACT TCC TTT AGG ATC TCA GGC TTC TGC AGT TCT CAT GAC TCC TAC TTT

        1035      1044      1053      1062      1071      1080
        TCA TCC TAG TCT AGC ATT CTG CAA CAT TTA TAT AGA CTG TTG AAA GGA GAA TTT

        1089      1098      1107      1116      1125      1134
        GAA AAA TGC ATA ATA ACT TCC ATC CCT GCT TAT TTT TAA TTT GGG AAA ATA

        1143      1152      1161      1170      1179      1188
        AAT ACA TTC GAA GGA AAA ACA AAA AGG GCG GCC CCC GAT TAT TGA GGG GTC
```

FIGURE 1C

```
      1197        1206          1215          1224             1233          1242
CCG AGC CCG AAC TCG TAA CCA TGT AAA ACC CGT TCC CCG GGG TAA AAT TGT AAT

      1251        1260          1269          1278             1287          1296
CCC CCC ACA ATT CCC CAA AAC ATA GGG CCC GGA AGC CTA AAG TTT AAA ACC CTG

      1305        1314          1323
GGG GGG CCT AAG GAG TTT ACC CAA ACT CCC TTT CT 3'
```

FIGURE 1D

```
         10          19          28          37          46      55
5'TCGA CGC CAT  TTT AAA GGG  ATT TAC TGC  ACG GAC TTC  TCC CAA GTT  CCT AGG

         64          73          82          91         100     109
CAT TAT CTT  CTG GAC CCT  ATC CTG CAG  AGG TGA AGC  GTC CCT TTG  GGG ACT CTC

        118         127         136         145         154     163
GCT GGG TGA  GAG GGA CAA  GAA ACA CCC  ACT AGG CCC  CAA CCC CGG  CAG CCA GCG

        172         181         190         199         208     217
GCT CGA GCA  TGC GCT GAG  AGT TTG TGC  AGC TGG CTG  CCG CCG CTG  CCT

        226         235         244         253         262     271
CGT CCG GAC  TCG GAG AGG  ACT TGG GAG  GGA CAG CGG  CGC TGG GAG  GTG GCT TAG

        280         289         298         307         316     325
CAG AGA CTT  TCC AGC AAC  TGC TGC CCA  GGA CTT TTT  TTT TTC TTT  TTC

        334         343         352         361         370     379
CCA GGC GGC  GAC GGC GGG  GGC AGA GGA  AGA GGA AGA  GAA AGC GTC  TCC

        388         397         406         415         424     433
AGC TGA AGC  CAA TGC AGC  CCT CCG GCT  CTC CGC GAA  GTT CCC TGC  CCC GAT

        442         451         460         469         478     487
GAG CCC CCG  TGC GTC CCC  GAC TAT CCC  CAG GCG GGC  GTG GGC CAC  CGG GCC
```

FIGURE 2A

```
        496         505         514         523         532         541
        CAG CGC CGA TCG CTG CCG TTT TGC CCT TGG GAG TAG GAT GTG GTG AAA GGA

        550         559         568         577         586         595
        TGG GGC TTC TCC CTT ACG GGG CTC ACA ATG GCC AGA GAA GAT TCC GTG AAG TGT
                                    M   A   R   E   D   S   V   K   C

        604         613         622         631         640         649
        CTG CGC TGC CTG CTC TAC GCC CTC AAT CTG CTC TTT TGG TTA ATG TCC ATC AGT
        L   R   C   L   L   Y   A   L   N   L   L   F   W   L   M   S   I   S

        658         667         676         685         694         703
        GTG TTG GCA GTT TCT GCT TGG ATG AGG GAC TAC CTA AAT AAT GTT CTC ACT TTA
        V   L   A   V   S   A   W   M   R   D   Y   L   N   N   V   L   T   L

        712         721         730         739         748         757
        ACT GCA GAA ACG AGG GTA GAG GAA GCA GTC ATT TTG ACT TAC TTT CCT GTG GTT
        T   A   E   T   R   V   E   E   A   V   I   L   T   Y   F   P   V   V

        766         775         784         793         802         811
        CAT CCG GTC ATG ATT GCT GTT TGC TGT TTC CTT ATC ATT GTG GGG ATG TTA GGA
        H   P   V   M   I   A   V   C   C   F   L   I   I   V   G   M   L   G

        820         829         838         847         856         865
        TAT TGT GGA ACG GTG AAA AGA AAT CTG TTG CTT GCA TGG TAC TTT GGA AGT
        Y   C   G   T   V   K   R   N   L   L   L   A   W   Y   F   G   S
```

FIGURE 2B

```
874         883         892         901         910         919
TTG CTT GTC ATT TTC TGT GTA GAA CTG GCT TGT GGC GTT TGG ACA TAT GAA CAG
 L   L   V   I   F   C   V   E   L   A   C   G   V   W   T   Y   E   Q

928         937         946         955         964         973
GAA CTT ATG GTT CCA GTA CAA TGG TCA GAT ATG ACT GTC ACT TTG AAA GCC AGG ATG
 E   L   M   V   P   V   Q   W   S   D   M   V   T   L   K   A   R   M

982         991         1000        1009        1018        1027
ACA AAT TAT GGA TTA CCT AGA TAT CGG TGG CTT ACT CAT GCT TGG AAT TTT TTT
 T   N   Y   G   L   P   R   Y   R   W   L   T   H   A   W   N   F   F

1036        1045        1054        1063        1072        1081
CAG AGA GAG TTT AAG TGC TGT GGA GTA TAT TTC ACT GAC TGG TTG GAA ATG
 Q   R   E   F   K   C   C   G   V   Y   F   T   D   W   L   E   M

1090        1099        1108        1117        1126        1135
ACA GAG ATG GAC TGG CCC CCA GAT TCC TGC TGT GTT AGA GAA TTC CCA GGA TGT
 T   E   M   D   W   P   P   D   S   C   C   V   R   E   F   P   G   C

1144        1153        1162        1171        1180        1189
TCC AAA CAG GCC CAC CAG GAA CTC AGT GAC CTT TAT CAA GAG GGT TGT GGG
 S   K   Q   A   H   Q   E   D   L   S   D   L   Y   Q   E   G   C   G

1198        1207        1216        1225        1234        1243
AAG AAA ATG TAT TCC TTT TTG AGA GGA ACC AAA CAA CTG CAG GTG CTG AGG TTT
 K   K   M   Y   S   F   L   R   G   T   K   Q   L   Q   V   L   R   F
```

FIGURE 2C

```
1252        1261        1270        1279        1288        1297
CTG GGA ATC TCC ATT GGG GTG ACA CAA ATC CTG GCC ATG ATT CTC ACC ATT ACT
 L   G   I   S   I   G   V   T   Q   I   L   A   M   I   L   T   I   T

1306        1315        1324        1333        1342        1351
CTG CTC TGG GCT CTG TAT TAT GAT AGA AGG GAG CCG GGG ACA GAC CAA ATG ATG
 L   L   W   A   L   Y   Y   D   R   R   E   P   G   T   D   Q   M   M

1360        1369        1378        1387        1396        1405
TCC TTG AAG AAT GAC AAC TCT CAG CAC CTG TCA TGT CCC TCA GTA GAA CTG TTG
 S   L   K   N   D   N   S   Q   H   L   S   C   P   S   V   E   L   L

1414        1423        1432        1441        1450        1459
AAA CCA AGC CTG TCA AGA ATC TTT GAA CAC ACA TCC ATG GCA AAC AGC TTT AAT
 K   P   S   L   S   R   I   F   E   H   T   S   M   A   N   S   F   N

1468        1477        1486        1495        1504        1513
ACA CAC TTT GAG ATG GAG GAG TTA TAA AAA GAA ATG TCA CAG AAG AAA ACC ACA
 T   H   F   E   M   E   E   L   *

1522        1531        1540        1549        1558        1567
AAC TTG TTT TAC TGG ACT TGT GAA TTT TTG AGT ACA TAC TAT GTG TTT CAG AAA

1576        1585        1594        1603        1612        1621
TAT GTA GAA ATA AAA ATG TTG CCA TAA AAT AAC ACC TAA GCA TAT ACT ATT CTA
```

FIGURE 2D

```
              1630.        1639        1648        1657        1666        1675
          TGC TTT AAA ATG AGG ATG GAA AAG TTT CAT GTC ATA AGT CAC CAC CTG GAC AAT

              1684        1693        1702        1711        1720        1729
          AAT TGA TGC CCT TAA AAT GCT GAA GAC AGA TGT CAT ACC CAC TGT GTA GCC TGT

              1738        1747        1756        1765        1774        1783
          GTA TGA CTT TTA CTG AAC ACA GTT ATG TTT TGA GGC AGC ATG GTT TGA TTA GCA

              1792        1801        1810        1819        1828        1837
          TTT CCG CAT CCA TGC AAA CGA GTC ACA TAT GGT GGG ACT GGA GCC ATA GTA AAG

              1846        1855        1864        1873        1882        1891
          GTT GAT TTA CTT CTA CCA ACT AGT ATA TAA AGT ACT AAT TAA ATG CTA ACA TAG

              1900        1909        1918        1927        1936        1945
          GAA GTT AGA AAA TAC TAA TAA CTT TTA TTA CTC AGC GAT CTA TTC TTC TGA TGC

              1954        1963        1972        1981        1990        1999
          TAA ATA AAT TAT ATA TCA GAA AAC TTT CAA TAT TGG TGA CTA CCT AAA TGT GAT

              2008        2017        2026        2035        2044        2053
          TTT TGC TGG TTA CTA AAA TAT TCT TAC CAC TTA AAA GAG CAA GCT AAC ACA TTG

              2062        2071        2080        2089        2098        2107
          TCT TAA GCT GAT CAG GGA TTT TTT GTA TAT AAG TCT GTG TTA AAT CTG TAT AAT
```

FIGURE 2E

```
        2116        2125        2134        2143        2152        2161
TCA GTC GAT TTC AGT TCT GAT AAT GTT AAG AAT AAC CAT TAT GAA AAG GAA AAT

        2170        2179        2188        2197        2206        2215
TTG TCC TGT ATA GCA TCA TTA TTT TTA GCC TTT CCT GTT AAT AAA GCT TTA CTA

        2224        2233        2242        2251        2260        2269
TTC TGT CCT GGG CTT ATA TTA CAC ATA TAA CTG TTA TTT AAA TAC TTA ACC ACT

        2278        2287        2296        2305        2314        2323
AAT TTT GAA AAT TAC CAG TGT GAT ACA TAG GAA TCA TTA TTC AGA ATG TAG TCT

        2332        2341        2350        2359        2368        2377
GGT CTT TAG GAA GTA TTA ATA AGA AAA TTT GCA CAT AAC TTA GTT GAT TCA GAA

        2386        2395        2404        2413        2422        2431
AGG ACT TGT ATG CTG TTT TTC TCC CAA ATG AAG ACT CTT TTT GAC ACT AAA CAC

        2440        2449        2458        2467        2476        2485
TTT TTA AAA AGC TTA TCT TTG CCT TCT CCA AAC AAG AAG CAA TAG TCT CCA AGT

        2494        2503        2512        2521        2530        2539
CAA TAT AAA TTC TAC AGA AAA TAG TGT TCT TTT TCT CCA GAA AAA TGC TTG TGA

        2548        2557        2566        2575        2584        2593
GAA TCA TTA AAA CAT GTG ACA ATT TAG AGA TTC TTT GTT TTA TTT CAC TGA TTA
```

FIGURE 2F

```
        2602     2611     2620     2629     2638     2647
        ATA TAC TGT GGC AAA TTA CAC AGA TTA AAT TTT ACA AGA GTA TAG TAT

        2656     2665     2674     2683     2692     2701
        ATT TAT TTG AAA TGG GAA AAG TGC ATT TTA CTG TAT TTT GTG TAT TTT GTT TAT

        2710     2719     2728     2737     2746     2755
        TTC TCA GAA TAT GGA AAG AAA ATT AAA ATG TGT CAA TAA ATA TTT TCT AGA GAG

        2764
        TAA AAA AAA AAA A 3'
```

FIGURE 2G

```
  1  MAKDNSTVRCFQGLLIFGNVIIGCCGIALTAECIF   2651154
  1  MAKDDSTVRCFQGLLIFGNVIIGMCSIALMAECIF   g443785

 36  FVSDQHSLYPLLEATDNDDIYGAAWIGIFVGICLF   2651154
 36  FVSDQNSLYPLLEATNNDDIYAAAWIGMSVGICLF   g443785

 71  CLSVLGIVGIMKSSRKILLAYFILMFIVYAFEVAS   2651154
 71  CLSVLGIVGIMKSNRKILLVYFILMFIVYAFEVAS   g443785

106  CITAATQRDFFTPNLFLKQMLERYQNNSPPNNDDQ   2651154
106  CITAATQRDFFTPNLFLKQMLERYQNNSPPNNDDQ   g443785

141  WKNNGVTKTWDRLMLQDNCCGVNGPSDWQKYTSAF   2651154
141  WKNNGVTKTWDRLMLQDNCCGVNGPSDWQKYTSAF   g443785

176  RTENNDADYPWPRQCCVMNNLKEPLNLEACKLGVP   2651154
176  RTENSDADYPWPRQCCVMNSLKEPLNLDACKLGVP   g443785

211  GFYHNQGCYELISGPMNRHAWGVAWFGFAILCWTF   2651154
211  GYYHSHGCYELISGPMNRHAWGVAWFGFAILCWTF   g443785

246  WVLLGTMFYWSRIEY   2651154
246  WVLLGTMFYWSRIDY   g443785
```

FIGURE 3

```
1    MAREDSVK-----CLRCLLYALNLLFWLMSISVL          2674553
1    MGEFNEKKATCGTVCLKYLLFTYNCCFWLAGLAVM         g2447007

30   AVSAWMRDYLNNVLTLTAETRVEEAVILTYFPVVH         2674553
36   AVGIWTLALKSDYISLLASS------TYLATAY           g2447007

65   PVMIAVCCFLIVGMLGYCGTVKRNLLLLAWYFGS          2674553
63   -ILVVAGVVVMVTGVLGCCATFKERRNLLRLYFIL         g2447007

100  LLVIFCVELACGVWTYEQELMVPVQWSDMVT-LKA         2674553
97   LLIFLLEIIAGILAY-----VYYQQLNTELKE            g2447007

134  RMTNYGLPRYRWLTH-----AWNFFQREFKCCGVV         2674553
125  NLKDTMVKRYHQSGHEGVSSAVDKLQQEFHCCGSN         g2447007

164  YFTDWLEMTEMDW-------PPDSCCVREFPG            2674553
160  NSQDW---QDSEWIRSGEADSRVVPDSCCKTMVAG         g2447007

189  CSKQAHQEDLSDLYQEGCGKKMYSFLRGTKQLQVL         2674553
192  CGKRDHASNIYKV-EGGCITKLETFIQE--HLRVI         g2447007

224  RFLGISIGVTQILAMILTITLLWALYYDRREPGTD         2674553
224  GAVGIGIACVQVFSMIFTCCLYRSL----------         g2447007
```

FIGURE 4A

```
259  Q M M S L K N D N S Q H L S C P S V E L L K P S L S R I F E H T S M A   2674553
249  - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -   g2447007

294  N S F N T H F E M E E L   2674553
249  - - - - - K L E H Y       g2447007
```

FIGURE 4B

```
  1  GCCC-TACCGTGT--GCGCAGAAAGAGAGGAGGCGCT    2651154dna
  1  GCGTGCAGAGCCGACACAGTACCAGGAGGAGGA        g443784

 33  TGCCTTCAGCTTGTGTGGGAAATCCCGAAGATGGCCA    2651154dna
 36  GGAG--AGGCTTGGGGGAAATCCTGAAGATGGCCA      g443784

 68  AAGACAACTCAACTGTTCGTTGCTTCCAGGGCCTG      2651154dna
 69  AAGACGACTCCACTGTTCGTTGCTTCCAGGGCCTG      g443784

103  CTGATTTTTGGAAATGTGATTATTGGTTGTTGCGG      2651154dna
104  CTGATTTTTGGAAATGTGATTATCGGTATGTGCAG      g443784

138  CATTGCCCCTGACTGCGGGAGTGCATCTTCTTTGTAT    2651154dna
139  CATCGCCCTGATGGCAGAGTGCATCTTCTTTGTAT      g443784

173  CTGACCAACACAGCCCTCTACCCACTGCTTGAAGCC     2651154dna
174  CAGACCAACACAAACAGCCCTCTACCCACTGCTTGAAGCC g443784

208  ACCGACAACGATGACATCTATGGGGCCTGCTGGAT      2651154dna
209  ACCAACAATGACGACATCTATGCGGCAGCCTGGAT      g443784

243  CGGCATATTTGTGGGGCATCTGCCTCTTCTGCCTGT     2651154dna
244  TGGCATGTCTGTTGGCATCTGCCTCTTCTGCCTCT      g443784
```

FIGURE 5A

```
278  CTGTTCTAGGCATTGTAGGCATCATGAAGTCCAGC   2651154dna
279  CTGTTCCTGGGCATCGTAGGCATCATGAAGTCCAAC  g443784

313  AGGAAAATTCTTCTGGCGTATTTCATTCTGATGTT   2651154dna
314  AGGAAAATTCTTCTGGTGTATTTCATCCTGATGTT   g443784

348  TATAGTATATGCCTTTGAAGTGGCATCTTGTATCA   2651154dna
349  TATTGTATATGCTTTTGAAGTGGCATCTTGTATCA   g443784

383  CAGCAGCAACACAACGAGACTTTTCACACCCAAC    2651154dna
384  CAGCAGCAACACAACGAGACTTTTCACACCCAAC    g443784

418  CTCTTCCTGAAGCAGATGCTAGAGAGGTACCAAAA   2651154dna
419  CTCTTCCTGAAGCAGATGCTGGAGAGATACCAAAA   g443784

453  CAACAGCCCTCCAAAACAATGATGACCAGTGGAAAA  2651154dna
454  CAACAGTCCTCCAAAACAATGATGACCAATGGAAAA  g443784

488  ACAATGGAGTCACCAAAACCCTGGGACAGGCTCATG  2651154dna
489  ACAATGGAGTCACCAAGACCCTGGGACAGGACTTATG g443784

523  CTCCAGGACAATTGCTGTGGCGTAAATGGTCCATC   2651154dna
524  CTCCAGGACAATTGCTGTGGTGTAAATGGCCCGTC   g443784
```

FIGURE 5B

```
558  A G A C T G G C A A A A T A C A C A T C T G C C T T C C G G A C T G    2651154dna
559  A G A C T G G C A G A A A T A C A C C C T C T G C C T T C C G G A C T G    g443784

593  A G A A T A A T G A T G C T G A C T A T C C C T G G C C C T C G T C A A    2651154dna
594  A G A A C G A T G C T G A C T A C C C C T G G C C C T C G T C A A    g443784

628  T G C T G T G T T A T G A A C A A T C T T A A A G A A C C C T C A A    2651154dna
629  T G C T G T G T T A T G A A C A G C C T T A A A G A A C C C T C A A    g443784

663  C C T G G A G G C T T G T A A A C T A G G C G T G C C T G G T T T T    2651154dna
664  C C T G G A C G C C T G C A A A T T A G G A G T G C C T G G A T A C T T    g443784

698  A T C A C A A T C A G G G C T G C T A T G A A C T G A T C T C T G G T    2651154dna
699  A C C A T A G T C A T G G C T G C T A T G A G C T G A T C T C T G G A    g443784

733  C C A A T G A A C C G A C A C G C C T G G G G G G T T G C C T G G T T    2651154dna
734  C C A A T G A A C C G A C A T G C C T G G G G G A G T T G C A T G G T T    g443784

768  T G G A T T T G C C A T T C T C T G C T G G A C T T T T T G G G T T C    2651154dna
769  T G G A T T T G C C A T T C T C T G T T T G G A C T T T C T G G G T T C    g443784

803  T C C T G G G T A C C A T G T T C T A C T G G A G C A G A A T T G A A    2651154dna
804  T C C T G G G T A C C A T G T T C T A C T G G A G C A G A A T T G A C    g443784
```

FIGURE 5C

TATTAAGCATAAAGTGTTGCCACCATACCCTTC 2651154dna
TATTAAGAAATGCACCATACCACTCCC g443784

CCCGAGTGACTCTGGATTTGGTGCTGGAACCAGCT 2651154dna
CAC-AGTGACTTGGATTTGGTGCTGGAAAT-GCT g443784

CTCTCCTAATATTCCACGTTTGTGCCCCAC---A 2651154dna
GTCTCCTAATGTTCTTACCTTTGTGCTGCCGGGAA g907

CTAACGTGTGTGTCTTACATTGCCAAGTCAGATGG 2651154dna
CTTACGCATTCCTTCCTACATTGCCAAGTACGTTGG g942

TACGGACTTCCTTTAGGATCTCAGGCTTCTGCAGT 2651154dna
TATGGGGTTCCTTTAAGCTCTCAGACT-CTGAAAT g977

TCTCATGACTCCTACTTTTCATCCTTAGTCTTAGCAT 2651154dna
TTTCAGCACATGTG-TTTTCACCCTGATCTAGGAT g1011

TCTGCAACATTTATATAGACTGTTGAAA--GGAGA 2651154dna
TCTGCAACATTGTTATAGACTGTAGGAAAGGGAGG g1045

ATTTGAAA-AATGCATAATAACTACTTCCATCCCT 2651154dna
ATTTAGGATAGTAGATAATAACTATTCCCATCTTT g1080

FIGURE 5D

```
1111  G C T T A T T T T T A A T T T G G G A A A A T A A A A T A C A T T C G A   2651154dna
1115  G T T T A T T T T T A A T G T G G G G G C A T A A A G A C A T T C C T     g443784

1146  A G G A A - - - - - - A A A C A A A A - - -                               2651154dna
1150  A G G A A C C T G T G T T A T A C T G C A A G T C T G T A T               g443784

1159  - - - - - - - - - - - - -                                                2651154dna
1185  T G G G A C A G C A A A T C T G C C C T G T A T T T C T C A C T G C T T   g443784

1159  - - A A A A G - - - G G C G G C C C C G A T T - - - A T                   2651154dna
1220  T C T A A A A G T A C C C T G A T G G C A C C C C A C T C C A G T A C     g443784

1180  T - - - - - - G A G G G G T C C -                                        2651154dna
1255  T C T T G C C T G G A A A A T C C C A T G G A C G G G A G C C T G         g443784

1190  - - - - - - - C G A G C                                                  2651154dna
1290  A T G G G C T G C C A G T C C A T G G G T C G C A A A G A G T C G G A C   g443784

1195  C C G A - - - A C T C G T A A C - - - C                                   2651154dna
1325  C C G A C T G G G C G A C T T C A C T T T C A C T T T C                   g443784

1209  A T G T A - - - A A - - - A C C C G T T C C - - -                         2651154dna
1360  A T G C A T T G G A G A A G G A A A T G G C A A C C C A C T C C A G T     g443784
```

FIGURE 5E

```
1225  ----------CCGGGG----GTTCTTGCCTGGAGAATCCCAGGGATGGAGGAGCC   2651154dna
1395                                                         g443784

1231  ------------TGGAGGGGCTGCCGTCTATGGGGTCACACAGAGTCG         2651154dna
1430                                                         g443784

1231  ----TAAA----GACACGACTGAAGCGACTTAGCAGCAGCAGCAGCA         2651154dna
1465                                                         g443784

1235  --ATTGTA----AAGGCTTTCATTGTATCAGTATTGTCCCAGTGAGA        2651154dna
1500                                                         g443784

1246  --------GAACTAAGGAGAAGACTGCTGAAACATCTTTTGAA            2651154dna
1535                                                         g443784

1246  --CCCAC----TTTGTTCTATGGTGGCTCCCACCTACAGACTCAAG          2651154dna
1570                                                         g443784

1251  -AATTCCCC-AAAAC----TGATTCTCTTAAAGCTAGCTTGGGAACCCCTTTATT  2651154dna
1605                                                         g443784

1264  ----ATAGGGCCCG----ATCCAAGACAAGGCCTGATCTTGAAACAAACAGTGG   2651154dna
1640                                                         g443784
```

FIGURE 5F

```
1274    - - G A A G C C T - - - - - - - - - - - - - - - - - - - - - - - - A A A G T - - - -    2651154dna
1675    T T G A A A T T T C C T C C T C A G A C A C T G C A G A G T A A T T C A                g443784

1286    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - T T A A A A - -        2651154dna
1710    T G C T G G T A A C C T C A A T T C T C C C A C T A A T T A A A A G T                  g443784

1292    - - - - C C C T - - - - - - - - - - - - - - - - G G G G G G C C T - - - -            2651154dna
1745    A C G T G A A C T T T T G G G A C A A A G G A G A G A C C T G T T A C              g443784

1306    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -            2651154dna
1780    A C A T T T A C C A C C T T C A A C C T A A A A C T G C T T T C C A A              g443784

1306    - - - - - - - - - - - - - - - - T T A C C - - - - - - - - - - -                      2651154dna
1815    C A G G G A A G A A A G C A A G C C A G C T G T T A C T T A G G T G A T            g443784

1318    - - - - - - - - - - - - - - - - - - - - - - C A A A - - - -                          2651154dna
1850    T T A G G G T G A T C T G T G C A C T G C A A A A T A T T T T C T T                  g443784

1322    - - - - C T C C C T T T C - - - - - - - - - - - - -                                  2651154dna
1885    C T G A T C T G T T T C C T T T G T G A T C C C T G A A G G A A T T T              g443784

1331    - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -                    2651154dna
1920    C T T A T A A C A A C A T T T G T C T T T A T A T A A A A T A A A A G A G            g443784
```

FIGURE 5G

2651154dna
g443784

1331
1955 AGTTTTAAAT

FIGURE 5H